

ABSTRACT OF THE DISCLOSURE

Concerning feature extraction of documents in text mining, a method and an apparatus for extracting features having the same nature as those by LSA are provided that require smaller memory space and simpler program and apparatus than the apparatus for executing LSA. Features of each document are extracted by feature extracting acts on the basis of a term-document matrix updated by term-document updating acts and of a basis vector, spanning a space of effective features, calculated by basis vector calculations. Execution of respective acts is repeated until a predetermined requirement given by a user is satisfied.

G:\DATA\WPDOCS3\DFW\DLT\TANI&ABE\14321.29PATENT.doc